JOURNAL OF CLINICAL MICROBIOLOGY, Sept. 1997, p. 2348–2358 Vol. 35, No. 9
0095-1137/97/$04.00+0

Parity among the Randomly Amplified Polymorphic DNA Method, Multilocus Enzyme Electrophoresis, and Southern Blot Hybridization with the Moderately Repetitive DNA Probe Ca3 for Fingerprinting *Candida albicans*

CLAUDE PUJOL,[1,2] SOPHIE JOLY,[1,2] SHAWN R. LOCKHART,[1] SEBASTIEN NOEL,[2] MICHEL TIBAYRENC,[2] AND DAVID R. SOLL[1]*

Department of Biological Sciences, University of Iowa, Iowa City, Iowa 52242,[1] and Laboratoire de Génétique Moléculaire des Parasites et des Vecteurs (UMR 9926, CNRS-ORSTOM), ORSTOM, BP 5045, 34032 Montpellier Cedex 1, France[2]

Received 11 March 1997/Returned for modification 29 May 1997/Accepted 5 June 1997

Randomly amplified polymorphic DNA (RAPD) analysis, multilocus enzyme electrophoresis (MLEE), and Southern blot hybridization with moderately repetitive DNA probes have emerged as effective fingerprinting methods for the infectious fungus *Candida albicans*. The three methods have been compared for their capacities to identify identical or highly related isolates, to cluster weakly related isolates, to discriminate between unrelated isolates, and to assess microevolution within a strain. By computing similarity coefficients between 29 isolates from three cities within the continental United States, strong concordance of the results is demonstrated for RAPD analysis, MLEE, and Southern blot hybridization with the moderately repetitive probe Ca3, and weaker concordance of the results is demonstrated for these three fingerprinting methods and Southern blot hybridization with the moderately repetitive probe CARE2. All methods were also demonstrated to be able to resolve microevolution within a strain, with the Ca3 probe exhibiting the greatest resolving power. The strong correlations demonstrated between polymorphic markers assessed by the four independent fingerprinting methods and the nonrandom association between loci demonstrated by RAPD analysis and MLEE provide evidence for strong linkage disequilibrium and a clonal population structure for *C. albicans*. In addition, a synapomorphic allele, *Pep-3A*, was found to be present in all members of one of the three clusters discriminated by RAPD analysis, MLEE, and Ca3 fingerprinting, supporting the concordance of the clustering capacities of the three methods, the robustness of the clusters, and the clonal nature of the clusters.

Candida albicans remains the most common and, in some cases, the most persistent fungal pathogen in humans, and this premier status is no doubt facilitated by its capacity to reside in the natural flora of healthy individuals as a commensal organism (36). Recently, fluconazole-resistant strains of *C. albicans* have emerged in compromised patients, especially those with AIDS-related disease (34, 68), underscoring the need for DNA fingerprinting methods that can be used to identify and monitor specific strains and assess the genetic relatedness of strains in broad epidemiological studies. Several recent observations together demonstrate that the epidemiology of *C. albicans* is complex. For instance, it has been demonstrated that strains of *C. albicans* can replace each other in recurrent infections (52, 58), that different body locations of the same healthy individual can harbor unrelated commensal strains (19, 26, 59), that the same body location can harbor different species or different strains of the same species (19, 35, 37, 60), that strains can be transferred from one individual to another (18, 26, 49, 52), that there is geographical localization of specific strains (49), and that colonizing strains can undergo microevolution and substrain shuffling in recurrent infections (25, 26, 52). These observations underscore the need for fingerprinting methods which provide measures of genetic distance between strains and which are amenable to computer-assisted methods for generating large databases for comparative and retrospective analyses (47, 52, 56).

Several methods have been used to fingerprint *C. albicans*, including electrophoretic karyotyping (2, 3, 5, 6, 12–14, 24, 27, 33, 67, 69), restriction fragment length polymorphism (RFLP) analysis (4, 28, 33, 66, 70), randomly amplified polymorphic DNA (RAPD) analysis (5, 6, 9, 17, 23, 24, 42, 51), Southern blot hybridization with a variety of moderately repetitive DNA probes (11, 20, 25, 29, 31, 46, 57), and multilocus enzyme electrophoresis (MLEE) (7, 8, 10, 21, 22, 38, 41). However, in most of these studies the patterns generated by the fingerprinting methods were not characterized for their level of discrimination, and the methods were not verified by other unrelated methods for their capacity to measure genetic distance between independent isolates. In addition, in some cases the reproducibility of the fingerprinting pattern was suspect and the amenability of the pattern to computer-assisted methods had not been demonstrated.

Of the several fingerprinting methods applied to *C. albicans*, Southern blot hybridization with the probe Ca3, RAPD analysis, and MLEE have emerged as being especially effective for different reasons. First, Southern blot hybridization with the midrepeat sequence Ca3 has proven to be reproducible and highly amenable to computer-assisted analysis (16, 25, 26, 47–50, 52, 59). Databases have been established for computer-assisted comparisons of the Ca3-generated patterns of strains from different studies and for retrospective analyses (19, 26, 52). The effectiveness of Ca3 in generating patterns which reflect genetic distance between independent isolates has re-

* Corresponding author. Mailing address: Department of Biological Sciences, 138 BB, University of Iowa, Iowa City, IA 52242. Phone: (319) 335-1117. Fax: (319) 335-2772.

TABLE 1. Origins of the 29 *C. albicans* isolates analyzed

Isolate	Origin (body location), geographical area of sampling	Date of isolation (mo/day/yr)
FC-1	Laboratory strain 3153A, smooth white phenotype	1961
FC-2	Laboratory strain 3153A, star phenotype	1961
FC-3	Laboratory strain WO-1 (systemic isolate), white phenotype, Iowa City	1984
FC-4	Laboratory strain WO-1, opaque phenotype	1984
FC-5	HIV-positive patient (oral isolate), El Paso, Tex.	12/1/1993
FC-6	HIV-positive patient (oral isolate), El Paso, Tex.	12/01/1993
FC-7	HIV-positive patient (oral isolate), El Paso, Tex.	12/01/1993
FC-8	HIV-positive patient (oral isolate), El Paso, Tex.	12/01/1993
FC-9	HIV-positive patient (oral isolate), El Paso, Tex.	12/1/1993
FC-10	Cancer patient (oral isolate), El Paso, Tex.	02/1/1994
FC-11	Vaginitis patient (oral isolate), Ann Arbor, Mich.	05/30/1990
FC-12	Same source as isolate FC-11 (fecal isolate), Ann Arbor, Mich.	06/06/1990
FC-13	Sexual partner of a vaginitis patient (seminal isolate), Ann Arbor, Mich.	11/24/1991
FC-14	Same source as FC-13 (oral isolate), Ann Arbor, Mich.	11/24/1991
FC-15	Vaginitis patient (vaginal isolate), Ann Arbor, Mich.	01/20/1992
FC-16	Vaginitis patient (vaginal isolate), Ann Arbor, Mich.	09/04/1992
FC-17	Vaginitis patient (vaginal isolate), Ann Arbor, Mich.	09/18/1992
FC-18	Sexual partner of patient from whom isolate FC-17 (fecal isolate) was obtained, Ann Arbor, Mich.	10/27/1992
FC-19	Vaginitis patient (vaginal isolate), Iowa City	03/15/1989
FC-20	Same source as FC-19 (vaginal isolate); Iowa City	07/12/1989
FC-21	Vaginitis patient (vaginal isolate), Iowa City	12/02/1993
FC-22	Vaginitis patient (vulval isolate), Iowa City	12/09/1993
FC-23	Vaginitis patient (vulval isolate), Iowa City	12/09/1993
FC-24	Sexual partner of patient from whom isolate FC-23 (oral isolate) was obtained, Iowa City	01/11/1994
FC-25	Healthy individual (oral isolate), Iowa City	02/07/1994
FC-26	Healthy individual (oral isolate), Iowa City	02/24/1994
FC-27	Healthy individual (oral isolate), Iowa City	02/08/1990
FC-28	Healthy individual (vulval isolate), Iowa City	02/15/1990
FC-29	Healthy individual (vulval isolate), Iowa City	02/07/1990

cently been verified through comparison with the unrelated CARE2 probe (26). The Ca3 probe has also been demonstrated to be highly effective in assessing microevolution within clonal populations over time, due primarily to the hypervariability of genomic sequences homologous to the C fragment of the probe (25, 26, 52).

The RAPD method of DNA fingerprinting, in which genomic DNA is amplified with a single short primer with an arbitrary sequence, has also gained favor because it is less time-consuming than Southern blot hybridization methods. However, this method has not been adequately verified by an alternative method to reflect genetic distances, nor has it been assessed for its capacity to monitor microevolution within a colonizing strain. Although MLEE is the most time-consuming method of the three, it should provide valid measures of genetic distance because it assesses defined multilocus differences. However, it may not be as sensitive in measuring rapid microevolutionary changes as the Ca3 probe and, thus, may not be an adequate method for revealing substrain shuffling or microevolution in an infecting population over time. MLEE is useful in assessing the mode of reproduction of microorganisms, and in a recent MLEE study of *C. albicans* strains in a population of human immunodeficiency virus (HIV)-positive individuals in France, the results indicated a clonal rather than sexual population structure (38), which was confirmed by Boerlin et al. (8) and Graser et al. (15).

It is therefore timely that the three potentially most effective methods for fingerprinting *C. albicans* be compared for their capacities to identify highly related isolates, to group or cluster weakly related isolates, to distinguish unrelated isolates, and to assess microevolution within a strain. Such a comparison should provide insight into the reliability of each method in measuring genetic distance and investigating the genetic diversity of *C. albicans*; and the data from the MLEE and RAPD analysis can be further used, through measurements of linkage disequilibrium between loci, to retest the clonal hypothesis of *C. albicans* reproduction (38). Linkage disequilibrium between the data sets obtained from the four independent fingerprinting methods (RAPD analysis, MLEE, Southern blot analysis with Ca3 and Southern blot hybridization with CARE2) was also assessed by computing the correlation between the distance matrices derived from each, an approach previously used in an analysis of the mode of reproduction of *Trypanosoma cruzi* (64).

MATERIALS AND METHODS

***C. albicans* strains.** Twenty-nine *C. albicans* isolates collected in the United States were analyzed. Each isolate was determined to be *C. albicans* with the ID 32C kit (bioMérieux, Marcy l'Etoile, France) for sugar assimilation and by hybridization with the species-specific probes Ca3 (43) and CARE2 (20). The health status of the host, the body location from which the isolate was obtained, and the geographical region of the host are listed for the relevant isolates in Table 1. The collection included switch phenotypes of strain 3153A (isolates FC-1 and FC-2) (53) and WO-1 (isolates FC-3 and FC-4) (54), multiple isolates from the same patients (isolates FC-11 and FC-12, FC-13 and FC-14, and FC-19 and FC-20), and isolates from vaginitis patients and their male sexual partners (isolates FC-17 and FC-18 and isolates FC-23 and FC-24). Isolates from the same individual, related individuals (e.g., sexual partners), or the same strain (i.e., switch phenotypes) were considered related in origin. Isolates from different individuals (excluding sexual partners) were considered unrelated in origin. The latter group represented 22 of the 29 isolates in the collection of isolates analyzed. Single colonies from each stock were reisolated and maintained on YPD (2% dextrose, 2% Bacto Peptone, 1% yeast extract, 2% agar) slants for experimental purposes.

Typing by RAPD analysis. A 5-ml suspension of cells (5×10^9 cells/ml) was mixed with an equal volume of glass beads (diameter, 0.45 mm) and disrupted by vigorous vortexing in CO_2 at 2°C. Sodium dodecyl sulfate (SDS) was added to a final concentration of 1% (wt/vol) and the DNA was purified by phenol-chloroform extraction (44). After precipitation in ethanol, the DNA was resuspended in Tris-EDTA buffer (pH 8.0). PCRs were performed in 0.5-ml microcentrifuge

tubes containing 25 μl of the following reaction mixture: 1 ng of *C. albicans* DNA; 2.5 μl of 10× buffer provided for *Taq* DNA polymerase; 0.5 U of *Taq* DNA polymerase (ATGC, Noisy-le-Grand, France); 200 μM (each) dATP, dCTP, dGTP and dTTP (Boehringer Mannheim, Meylan, France); and 0.2 μM one of the primers listed below. The reaction mixture was overlaid with one drop of mineral oil, and amplification was performed in a CROCODILE II thermal cycler (Appligene, Illkirch, France) programmed for 45 cycles of 1 min at 94°C, 1 min at 36°C, and 2 min at 73°C. The amplification products were separated by electrophoresis in a 1.3% agarose gel and were detected by staining with ethidium bromide. Data were obtained for the following eight primers: OPE-03 (CCAGATGCAC), OPE-04 (GTGACATGCC), OPE-12 (TTATCGCCCC), OPE-18 (GGACTGCAGA), OPE-19 (ACGGCGTATG), OPE-20 (AACG GTGACC), OPF-10 (GGAAGCTTGG), and OPF-12 (ACGGTACCAG) (Operon Technologies, Alameda, Calif.).

Typing by MLEE. Cytosolic extraction, starch gel electrophoresis, and enzymatic assays were performed by methods described previously for 21 enzyme loci (39). The activities of the following enzymes (with Enzyme Commission numbers) were analyzed: malic dehydrogenase (EC 1.1.1.37), glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49), sorbitol dehydrogenase (EC 1.1.1.14), isocitrate dehydrogenase (EC 1.1.1.42), alcohol dehydrogenase (EC 1.1.1.1), superoxide dismutase (EC 1.15.1.1), hexokinase (EC 2.7.1.1), pyruvate kinase (EC 2.7.1.40), aspartate aminotransferase (EC 2.6.1.1), phosphoglucomutase (EC 5.4.2.2), esterase (EC 3.1.1.1), leucine aminopeptidase (EC 3.4.11.1), peptidase 1 (EC 3.4.13.18; substrate, Val-Leu), peptidase 2 (EC 3.4.11.4; substrate, Leu-Gly-Gly), peptidase 3 (EC 3.4.13.9; substrate, Phe-Pro), aldolase (EC 4.1.2.13), fumarase (EC 4.2.1.2), mannose-6-phosphate isomerase (EC 5.3.1.8), and glucose-6-phosphate isomerase (EC 5.3.1.9). Malate dehydrogenase and hexokinase enzymatic activities were each expressed by two genetically interpretable loci: by *Mdh-1* and *Mdh-2* and by *Hk-1* and *Hk-2*, respectively.

DNA fingerprinting with midrepetitive species-specific probes (Ca3 and CARE2). Southern blot hybridization was performed as described previously (47). Briefly, cells from stored agar slants were transferred to YPD agar plates and harvested at the early stationary phase for DNA extraction. The extraction method was conducted as described by Scherer and Stevens (45). The DNA of each isolate was then digested with *Eco*RI (4 U/μg of DNA) for 16 h at 37°C. Digested DNA (3 μg/lane) was electrophoresed in a 0.8% agarose gel overnight at 35 V. The gel was stained with ethidium bromide to assess loading and was transferred by capillary blotting to a nitrocellulose membrane. The membrane was sequentially hybridized with the Ca3 probe (1, 43, 57) and the CARE2 probe (20, 26) as described previously (26, 49, 57). The CARE2 probe was a generous gift from Brent Lasker of the Centers for Disease Control and Prevention, Atlanta, Ga. In sequential hybridization, the membrane was first prehybridized (7 h at 65°C) with 100 μg of denatured calf thymus DNA per ml and was then hybridized with random primer-labeled Ca3 probe (overnight at 65°C). These steps were performed in a solution containing 5× SSPE (1× SSPE contains 10 mM NaH_2PO_4 [pH 7.5], 10 nM EDTA, and 0.18 M NaCl), 5% dextran sulfate, and 0.3% SDS. The membrane was washed at 45°C with a solution of 2× SSPE containing 0.2% SDS and was then exposed to XAR-S film (Eastman Kodak Co., Rochester, N.Y.) with a Cronex Lightning-Plus intensifying screen (Du Pont Co., Wilmington, Del.). The blot was stripped of the Ca3 probe by heating to 80°C for 15 min in 1 mM EDTA, rehybridized with random primer-labeled CARE2 probe, and again exposed to XAR-S film.

Cluster analyses. Dendrograms were generated for the different fingerprinting methods by using the same measurement of relatedness, a similarity coefficient (S_{AB}) based on band positions computed with the Dendron software package, version 2.0 (Solltech, Iowa City, Iowa). This S_{AB} measures the proportion of bands with the same molecular weights in the patterns of two isolates by the following formula: $S_{AB} = 2E/(2E + a + b)$, where E is the number of bands shared by strains A and B, a is the number of bands unique to strain A, and b is the number of bands unique to strain B. For the present study, an S_{AB} of 1.00 represents identically matched bands (i.e., all bands in the patterns of isolates A and B match), an S_{AB} of 0.0 represents no matches, and S_{AB}s ranging from 0.01 to 0.99 represent increasing proportions of matched bands. Dendrograms based on S_{AB} values were generated by the unweighted pair-group method (55).

For computing S_{AB}s between pairs of isolates with the Ca3 or CARE2 probes, autoradiograms of the Southern blot hybridization patterns were digitized into the Dendron data file with a Scanjet IIcx flatbed scanner (Hewlett-Packard Co., Palo Alto, Calif.) equipped with a transparency option. Distortions in the gel were straightened (unwarping option of Dendron), and each lane was automatically identified and scanned. The bands were automatically identified and then verified and edited manually, and matrices were generated on the basis of band positions (molecular weights). From this information, intermediary band data files that contained the genetic makeup for each strain (i.e., the presence or absence of bands) were generated. Using these band data files, S_{AB}s were then computed between each pair of isolates and dendrograms based on S_{AB}s were generated.

For computing S_{AB}s between pairs of isolates by MLEE and RAPD analysis, the patterns for each isolate were entered into band data files, and S_{AB}s were computed on the basis of the proportion of matches. In the case of MLEE, alleles were used in the matching procedure. For generating a combined dendrogram based on the fingerprinting patterns obtained with Ca3 and CARE2, the information contained in the different band data files were combined in a new band data file which was then processed as described above.

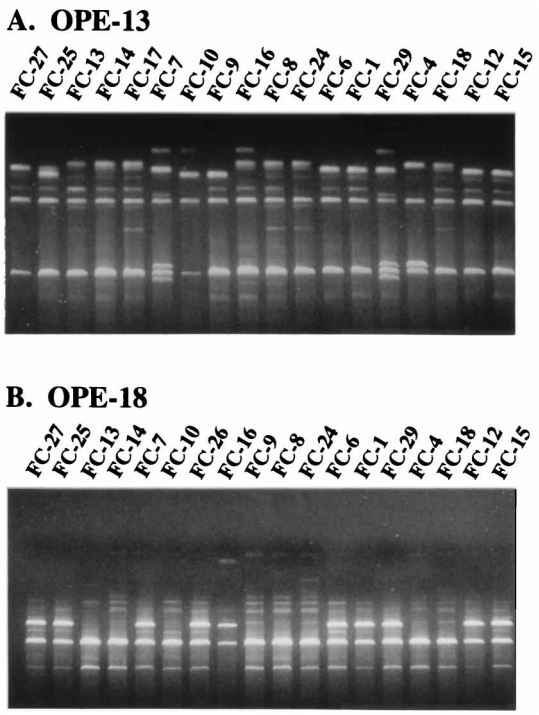

FIG. 1. RAPD patterns obtained with the primers OPE-3 (A) and OPE-18 (B) for different *C. albicans* isolates. Total genomic DNA from each isolate was amplified by PCR. Amplification products were analyzed by electrophoresis in 1.3% agarose gels and were detected by staining with ethidium bromide.

Comparison of methods by a nonparametric Mantel test. In order to compare the four fingerprinting methods for their capacities to measure genetic distance between isolates, correlations were estimated between every possible pair of S_{AB} matrices (i.e., RAPD analysis versus MLEE, RAPD analysis versus hybridization with Ca3, etc.) by the nonparametric Mantel test (30). In the first step, the correlation between two matrices is computed by using the coefficient of Pearson. Then, the entries of the matrix from one method were randomly permuted (10^4 iterations generated by Monte Carlo simulations), and correlations were made between the unmodified and the modified matrices in each case. This gives a distribution of Pearson coefficients. If less than 5% of the iterations (500 in this case, since there were 10^4 iterations) show a Pearson coefficient higher than or equal to the coefficient computed without permutation, the correlation is statistically positive (30).

Statistical tests for linkage disequilibrium. The four statistical tests used have been described in detail in previous publications (62, 63, 65). These tests are independent of the assumption of ploidy level, and the levels of significance for nonrandom association between loci were computed under the null hypothesis of random recombination (panmixia). The definitions of the four computed probabilities (*d1*, *d2*, *e*, and *f*) are given in Results.

RESULTS

RAPD analysis. Each of the 29 isolates in this study (Table 1) was analyzed with eight individual primers, each of 10 bp, by the RAPD method. These were selected from 40 primers tested for their capacities to discriminate variability and their reproducibilities. Examples of the patterns obtained with two primers with 18 isolates are presented in Fig. 1. All amplifications were repeated at least twice. In each case, the most intense bands in the patterns were reproducible, even with different extracts of DNA from the same isolate. However, differences were observed in some low-intensity bands. Therefore, in comparative analyses, only high-intensity bands were used. Each primer generated between one and six major bands for an individual isolate, with an average of 2.75 bands for the

TABLE 2. Polymorphic characteristics of bands analyzed by each of the four fingerprinting methods

Fingerprinting method	Total no. of bands analyzed[a]	No. of polymorphic bands (%)[b]	No. of polymorphic bands per isolate		Avg S_{AB}[c]
			Range	Avg	
RAPD analysis	31	16 (51.6)	4–10	7	0.87 ± 0.06
MLEE	45	31 (68.9)	7–13	9.6	0.82 ± 0.07
Analysis with Ca3	44	39 (88.6)	6–16	11.2	0.65 ± 0.11
Analysis with CARE2	48	48 (100)	12–24	17.2	0.50 ± 0.13

[a] In the case of RAPD analysis this represented the sum of the major bands obtained in the eight individual amplifications for the 29 isolates; in the case of MLEE, this represented the sum of the bands obtained for the 21 enzyme loci examined in the 29 isolates; in the cases of Ca3 and CARE2, this represented all bands with different molecular weights observed in the 29 analyzed isolates.

[b] A polymorphic band varied in at least one isolate.

[c] The average S_{AB}s were calculated for the 231 possible pairwise comparisons (S_{AB}s) among the 22 *C. albicans* isolates from unrelated hosts.

29 isolates analyzed. The total number of bands generated by the eight primers was 31 (Table 2). Only 16 were polymorphic and an average of 7 were polymorphic per isolate (Table 2). Among the 29 isolates analyzed, 17 different RAPD profiles (the combined results of amplifications with the eight different primers) were observed. A dendrogram of relatedness for the 29 analyzed isolates based on S_{AB}s computed from band positions is presented in Fig. 2A. The average S_{AB} for the 29 isolates analyzed was 0.87 ± 0.06. This value was highly similar to the average S_{AB} computed for only the 22 unrelated isolates (Table 2). The lowest connecting S_{AB} in the dendrogram was 0.82 (Fig. 2A). The RAPD analysis grouped all known related isolates (Table 1) with similarity coefficients of between 0.98 and 1.00 (Table 3). It also assessed as identical several isolates from unrelated hosts (Table 3). By using the average S_{AB} of 0.87 as a threshold, RAPD analysis generated three general clusters in a dendrogram (Fig. 2A). The three clusters included all but 1 of the 29 analyzed isolates (isolate FC-10).

MLEE. Each of the 29 isolates in this study (Table 1) was analyzed by MLEE. Of the 21 loci analyzed, 13 (62%) exhibited variability between two or more isolates, yielding 23 different enzyme profiles for the 29 analyzed isolates. Examples of two polymorphic enzyme activities (mannose-6-phosphate isomerase and hexokinase) are presented in Fig. 3. In the case of mannose-6-phosphate isomerase, a single locus (*Mpi*) which included five alleles among the 29 isolates analyzed was observed (four alleles are shown in Fig. 3A). In the case of hexokinase, two loci were observed: *Hk-1*, which included three alleles, and *Hk-2*, which included two alleles (Fig. 3B). The total number of bands generated by MLEE was 45 (Table 2). Only 31 were polymorphic and an average of 9.6 were polymorphic per isolate (Table 2). The average number of alleles for the variable loci in the 29 isolates analyzed was 2.6 per locus and ranged between 2 and 5. The eight monomorphic loci were *Sdh*, *Idh*, *Sod*, *Aat*, *Pgm*, *Est*, *Ald*, and *Fum*. The results of the individual MLEE analyses were pooled for each isolate, and a dendrogram of relatedness of the 29 analyzed isolates based on the S_{AB}s computed from the pooled data is

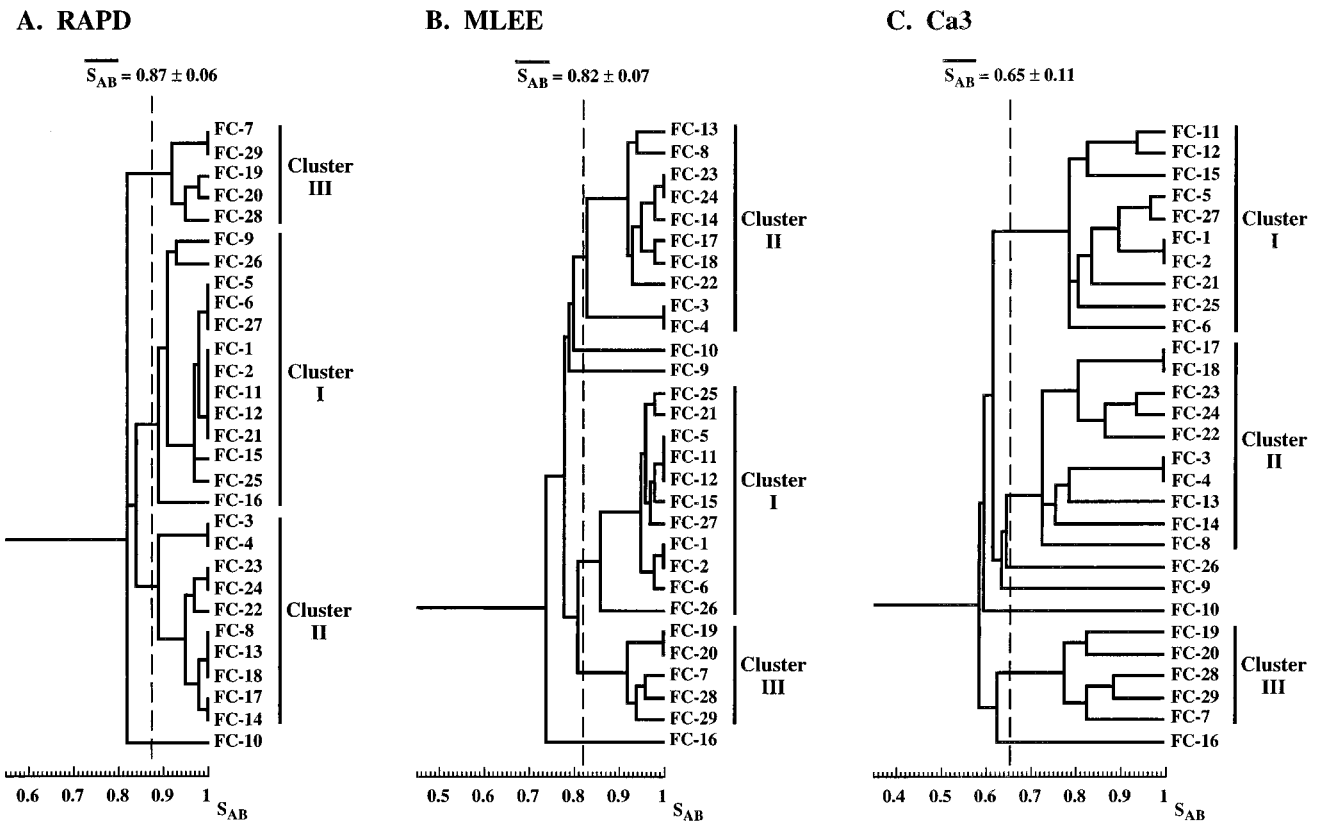

FIG. 2. Dendrograms based on the computed S_{AB}s between the 29 isolates fingerprinted by RAPD analysis (A), MLEE (B), and Southern blot hybridization with the DNA probe Ca3 (C). In each case the average S_{AB} is indicated as a dashed line. Note that each dendrogram contains three similar clusters of isolates (clusters I, II, and III).

TABLE 3. Comparison of the discriminatory abilities of the different fingerprinting methods in assessing identical or highly related isolates[a]

Isolate origin and group	Isolate similarity			
	RAPD	MLEE	Ca3	CARE2
Unrelated origin				
FC-7 = FC-29		FC-5 = FC-11/12		
FC-5 = FC-6 = FC-27				
FC-1 and FC-2 = FC-11 and FC-12 = FC-21				
FC-8 = FC-13 = FC-18		(FC-23 and FC-24; FC-14)		
FC-14 = FC-17		(FC-21 ≈ FC-25)	(FC-5 ≈ FC-27)	
(FC-5, FC-6, FC-27 ≈ FC-1 and FC-2, FC-11 and FC-12, FC-21)		(FC-5, FC-11 and FC-12 ≈ FC-15; FC-27)		
(FC-8, FC-13, FC-18 ≈ FC-17, FC-14)		(FC-1 and FC-2 ≈ FC-6)		
Related origin				
FC1 vs FC-2	Id	Id	Id	Id
FC-3 vs FC-4	Id	Id	Id	Id
FC-11 vs FC-12	Id	Id	HR (2)	HR (2)
FC-13 vs FC-14	HR (1)	MR (3)	MR (8)	MR (12)
FC-17 vs FC-18	HR (1)	HR (1)	Id	Id
FC-19 vs FC-20	HR (1)	Id	MR (7)	HR (1)
FC-23 vs VC-24	Id	Id	HR (2)	HR (1)

[a] The S_{AB} *of identical (equals signs or Id) isolates is 1.00 for RAPD and lysis, MLEE, and analysis with Ca3 and CARE2. The* S_{AB} thresholds of highly related (HR) isolates for the four methods are as follows: RAPD analysis, 0.98; MLEE, 0.98; analysis with Ca3, 0.94; analysis with CARE2, 0.94; S_{AB} thresholds for moderately related (MR) isolates for the three relevant methods are as follows: MLEE, 0.92; analysis with Ca3, 0.76; analysis with CARE2, 0.64; Data for analysis with CARE2 were not represented in the section with data for isolates of unrelated origin because no isolates of unrelated origin displayed identical or highly related genotypes. For isolates with unrelated origins, parentheses indicate that the isolates are highly related. For isolates with related origins, parentheses indicate the number of band differences. The symbol "≈" signifies "highly related but nonidentical."

presented in Fig. 2B. The average S_{AB} for all isolates analyzed was 0.83 ± 0.07, and the average S_{AB} computed for the 22 unrelated isolates was 0.82 ± 0.07. The lowest connecting S_{AB} in the dendrogram was 0.73 (Fig. 2B). MLEE grouped most of the known related pairs of isolates (Table 1) with S_{AB}s of between 0.98 and 1.00, as did the RAPD analysis, but it grouped isolates FC-13 and FC-14 with an S_{AB} of 0.92 (the RAPD analysis grouped this pair with an S_{AB} of 0.98) (Table 3). The MLEE analysis also grouped as identical the apparently unrelated isolate FC-5 and isolates FC-11 and FC-12 and as highly related several additional, apparently unrelated isolates (Table 3).

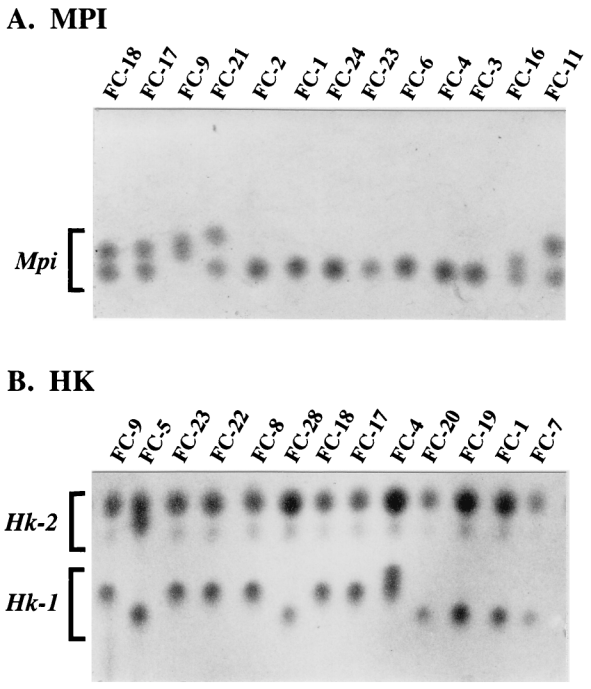

FIG. 3. Starch gel electrophoresis showing the enzyme phenotypes for mannose-6-phosphate isomerase (MPI) (A) and hexokinase (HK) (B) in 13 *C. albicans* isolates. Total cytosolic extracts were electrophoresed in starch gels under nondenaturating conditions. The enzyme activities were visualized by specific enzyme staining procedures. While the mannose-6-phosphate isomerase enzyme activity was expressed by a unique locus (*Mpi*), the hexokinase activity was expressed by two loci (*Hk-1* and *Hk-2*).

MLEE generated three general clusters at the average S_{AB} threshold of 0.82, which included all but 3 of the 29 isolates analyzed (isolates FC-9, FC-10, and FC-16) (Fig. 2B). While 100% of the isolates separated into the three clusters by the MLEE method were separated into the respective clusters by the RAPD method, only 90% of the isolates separated into the three clusters by the RAPD method were in the comparative clusters separated by MLEE. Clusters II and III of the dendrogram generated by MLEE contained the same isolates as clusters II and III of the dendrogram generated by RAPD analysis. Isolates FC-3 and FC-4, which barely grouped in cluster II of the dendrogram generated by RAPD analysis, barely grouped in cluster II of the dendrogram generated by MLEE. Similarly, isolate FC-26 barely grouped in cluster I of the dendrograms generated by RAPD analysis and MLEE. In contrast, isolates FC-9 and FC-16, which barely grouped in cluster I of the dendrogram generated by RAPD analysis, were not associated with cluster I of the dendrogram generated by MLEE.

It should be noted that allele *Pep-3A* of the *Pep-3* locus was found in all isolates in cluster II but in no other isolates, including all isolates in clusters I and III. Therefore, allele *Pep-3A* of *Pep-3* represents a synapomorphic marker (64), and its unique presence in cluster II suggests that all of the isolates in this cluster may have originated from a common progenitor.

Southern blot hybridization with Ca3. Each of the 29 isolates in this study (Table 1) was fingerprinted with the DNA probe Ca3. Ca3 generated patterns of between 11 and 21 bands (Table 2) in the range of 2.0 to >7.9 kb for each isolate,

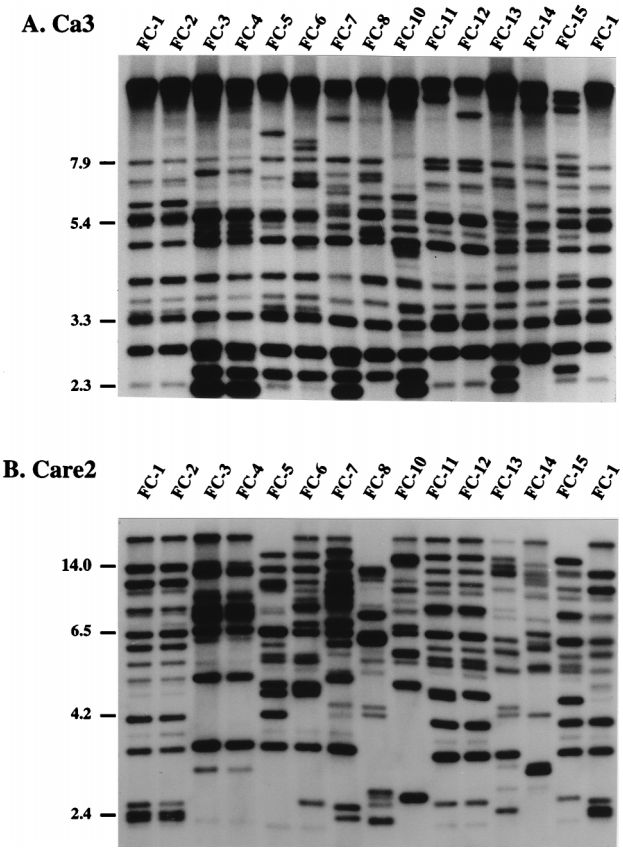

FIG. 4. Hybridization patterns of isolates FC-1 through FC-8 and FC-10 through FC-15 obtained with Ca3 (A) and CARE2 (B). Total genomic DNA from each isolate was digested with *Eco*RI, and Southern blots were sequentially hybridized with the Ca3 and CARE2 probes. Molecular sizes are presented in kilobases to the left of each hybridization pattern.

including invariant, moderately variant, and hypervariable bands (1, 47). Examples of the patterns obtained by hybridization with Ca3 for isolates FC-1 through FC-8 and FC-10 through FC-15 are presented in Fig. 4A. The reproducibilities of the patterns for both intense and faint bands are evident in the repeat patterns for isolate FC-1 in the outermost lanes, isolate FC-1 versus isolate FC-2 (switch phenotypes of strain 3153A), and isolate FC-3 versus isolate FC-4 (switch phenotypes of strain WO-1). Minor differences in highly related isolates are evident in the lanes for isolate FC-11 versus isolate FC-12 (isolates from the same individual), in which a single change in a high-molecular-weight band is evident. The total number of bands generated by Ca3 was 44 (Table 2). Thirty-nine were polymorphic, and an average of 11.2 were polymorphic per isolate (Table 2). Among the 29 isolates analyzed, 26 different Ca3 hybridization profiles were observed. A dendrogram based on band positions alone in each of the patterns of the 29 isolates analyzed is presented in Fig. 2C. The average S_{AB} for all isolates analyzed was 0.66 ± 0.11, and the average S_{AB} for the 22 isolates from unrelated hosts was 0.65 ± 0.11 (Table 2). These are well below the comparable average S_{AB}s computed for both RAPD analysis and MLEE (Table 2) and reflects a more complex pattern with increased variability. The lowest S_{AB} in the dendrogram generated by hybridization with Ca3 was 0.59, which is again well below the lowest S_{AB} for either the dendrogram generated by either RAPD analysis or MLEE (Fig. 2). The analysis with Ca3 grouped three pairs of known related isolates as identical: isolates FC-1 and FC-2, FC-3 and FC-4, and FC-17 and FC-18. These three pairs of isolates were also grouped in the S_{AB} range of 0.98 to 1.00 by the RAPD and MLEE methods (Table 3). The analysis with Ca3 also grouped the pairs of isolates FC-11 and FC-12 and isolates FC-23 and FC-24 as highly related, with S_{AB}s of 0.94. However, the pairs of isolates FC-13 and FC-14 and isolates FC-19 and FC-20, which were found to be highly similar or identical by RAPD analysis and MLEE, were grouped, but at reduced S_{AB}s (Table 3). The analysis with Ca3 also grouped one pair of unrelated isolates, isolates FC-5 and FC-27, which were also grouped by the RAPD and MLEE methods at high S_{AB} values, but the level of grouping of apparently unrelated isolates at high S_{AB}s was greater by RAPD analysis and MLEE than by the method with Ca3 (Table 3). The analysis with Ca3 generated three general clusters at the average S_{AB} threshold of 0.65, which included all but 4 isolates of the 29 isolates analyzed in this study (isolates FC-9, FC-10, FC-16, and FC-26) (Fig. 2C). Cluster I in the dendrogram generated by analysis with Ca3 contained 10 isolates present in cluster I of the dendrogram generated by MLEE but lacked isolate FC-26, which was barely grouped in cluster I of the dendrogram generated by MLEE and also lacked the same two isolates (isolate FC-9 and FC-16) which were present in cluster I of the dendrogram generated by RAPD analysis but not in cluster I of the dendrogram generated by MLEE. Clusters II and III in the dendrogram generated by analysis with Ca3 contained the same isolates as clusters II and III of the dendrograms generated by RAPD analysis and MLEE (Fig. 2).

Southern blot hybridization with CARE2. Each of the 29 isolates in this study (Table 1) was also fingerprinted with the moderately repetitive DNA sequence CARE2 (20), which has been demonstrated to be unrelated to Ca3 (20, 26). CARE2 generated banding patterns of between 12 and 24 bands (Table 2) in the range of 2 to >14 kb for each isolate. Examples of the patterns for isolates FC-1 through FC-8 and isolates FC-10 through FC-15 are presented in Fig. 4B. As in the case of Ca3, the reproducibility of patterns is evident for isolate FC-1 in the outer two lanes, for isolates FC-1 versus isolate FC-2, and for isolate FC-3 versus isolate FC-4. Minor differences in highly related isolates are also evident in the lanes for isolate FC-11 versus isolate FC-12. A dendrogram based on band positions alone is presented in Fig. 5A. The average S_{AB} for all isolates analyzed was 0.51 ± 0.14, and the average S_{AB} for the 22 isolates from unrelated hosts was 0.50 ± 0.13. The CARE2 analysis generated two general clusters at the average S_{AB} threshold of 0.50 (Fig. 5A) which fairly correlated with clusters I and II in the dendrograms generated by RAPD analysis, MLEE and the analysis with Ca3 (Fig. 2). Since in the dendrogram generated by analysis with CARE2 (Fig. 5A) there was fragmentation of two of the clusters (clusters II and III) discriminated in concert by RAPD analysis, MLEE, and analysis with Ca3, we have noted in the dendrogram generated by analysis with CARE2 the members of the clusters separated by the method with Ca3. Cluster I in the dendrogram generated by analysis with CARE2 (Fig. 5A) contained all but one isolate (FC-6) in cluster I of the dendrogram generated by analysis with Ca3 (Fig. 2C). Cluster II of the dendrogram generated by analysis with CARE2 (Fig. 5A) also contained 7 of the 10 isolates in cluster II of the dendrogram generated by analysis with Ca3 (Fig. 2C) but was missing isolates FC-3, FC-4, and FC-8 (Fig. 5A). As noted above, FC-3 and FC-4 were only weakly associated with cluster II in the RAPD analysis (Fig. 2A) and MLEE (Fig. 2B). CARE2 grouped only three of the five isolates in cluster III of the dendrograms generated by RAPD analysis, MLEE, and analysis with Ca3 (Fig. 5A).

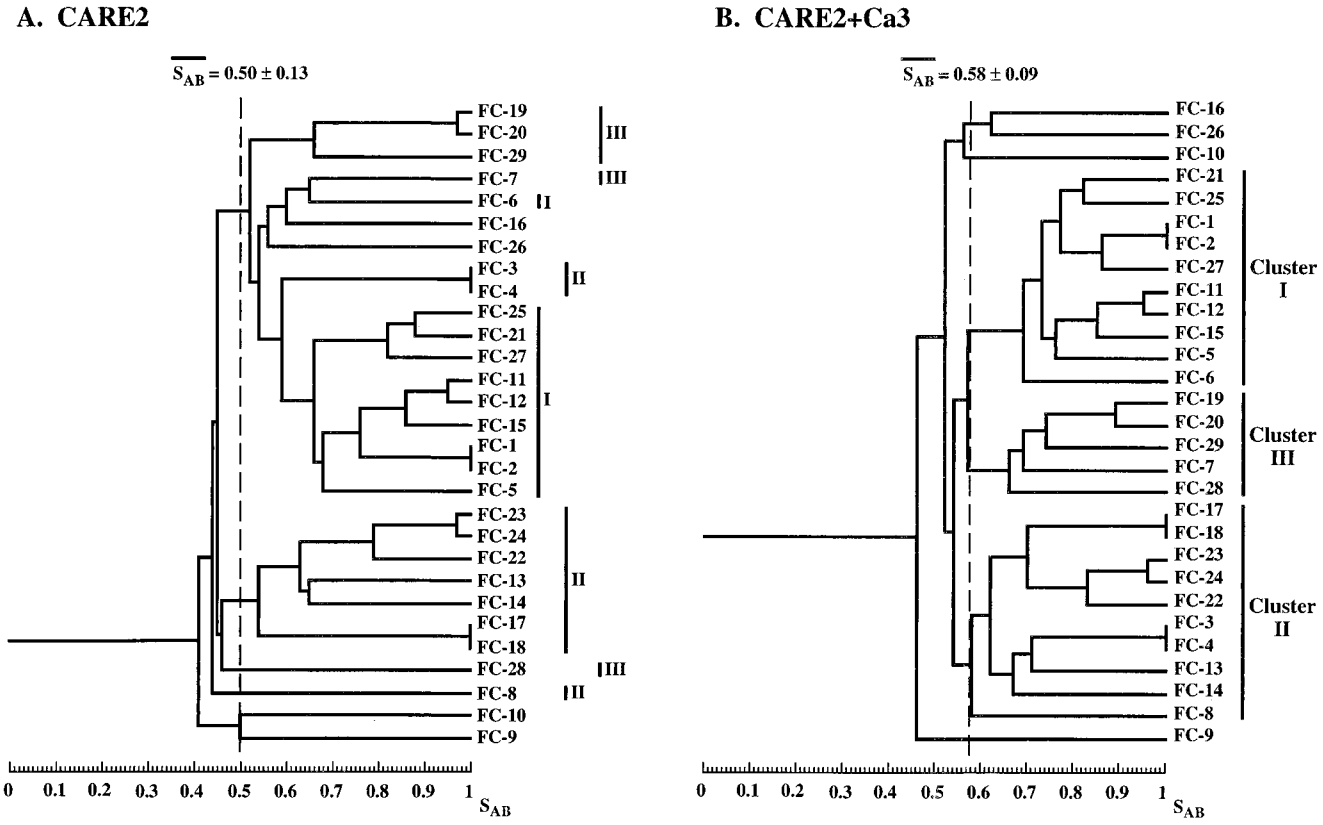

FIG. 5. Dendrograms based on the computed S_{AB}s between the 29 isolates fingerprinted by hybridization with CARE2 (A) and by combining the banding data from analyses with both CARE2 and Ca3 (B).

The strong congruence in clustering between RAPD analysis, MLEE, and analysis with Ca3 and the weaker congruence for the CARE2 method suggest that the patterns generated by CARE2 include either homoplasy (reversion or conversion of pattern) or a differential increase in the rate of randomization of patterns. To examine these possibilities, we generated a dendrogram combining the banding data from analyses with both Ca3 and CARE2 (Fig. 5B). In the combined dendrogram, clusters I, II, and III reemerged with the same memberships as in the dendrogram obtained with data from the analysis with Ca3 alone (Fig. 2C). This result suggests that the patterns obtained by analysis with CARE2 display either a considerable level of homoplasy or a more rapid rate of random reorganization, especially for the cluster III isolates, and that the structuring information present in the patterns obtained by analysis with Ca3 is strong enough to reintroduce order. The data strongly suggest that the patterns obtained by analysis with CARE2 display overall a more rapid rate of divergence by random reorganization than those obtained by analysis with Ca3. This is reflected by the number of polymorphic bands analyzed by both methods (48 for analysis with CARE2 compared to 39 for analysis with Ca3) (Table 2) and in the lower average S_{AB} computed for analysis with CARE2 (0.50 ± 0.13 compared to 0.65 ± 0.11 for analysis with Ca3) (Table 2).

Discriminatory powers of the fingerprinting methods tested. The fingerprinting methods compared in this study exhibit two characteristics which are important in epidemiological studies, the capacity to group apparently unrelated isolates (e.g., from unrelated individuals in the same geographical location or from individuals in different geographical locales) and the capacity to discriminate differences in highly related isolates (e.g., isolates from recurrent infections in the same individual or from sexual partners). One would expect grouping capabilities to depend on the number of variable or polymorphic bands in a pattern, which would presumably provide increased levels of discrimination and comparison. In fact, the average number of polymorphic bands per strain for RAPD analysis, MLEE, and analyses with the Ca3 and CARE2 probes were 7.0, 9.6, 11.2, and 17.2, respectively (Table 2). However, even though CARE2 generated the largest number of polymorphic bands, and, therefore, the lowest average S_{AB} in a dendrogram (Table 2), it still resulted in the poorest level of clustering. The most noteworthy aspect of these results is that in spite of the significant differences in the proportion and absolute number of polymorphic bands, the clustering capacities of RAPD analysis, MLEE, and analysis with Ca3 were similar (Fig. 2).

In order to assess statistically the correlations between the four data sets, we used the Mantel nonparametric test to compare the nonrandom association between each pair of the four similarity coefficient matrices obtained for the 22 isolates that were unrelated in origin. All of the Mantel tests were significant at the 5% level of significance, suggesting high levels of correlation between all four methods. However, differences existed in the correlations between analysis with Ca3, MLEE, and RAPD analysis versus the correlations between analysis with CARE2 and either analysis with Ca3 or RAPD analysis. While the probabilities between the former three methods were all below 10^{-4}, the probability of correlation between analysis with CARE2 and RAPD analysis was 3.5×10^{-3} and that between analysis with CARE2 and analysis with Ca3 was

2×10^{-2}. This result again indicates a weaker correlation between analysis with CARE2 and two of the three other methods and again demonstrates that when compared to the three other fingerprinting methods, analysis with CARE2 is less congruent. The nonrandom association demonstrated by the Mantel test between the four sets of independent polymorphic markers obtained by the four fingerprinting systems also indicates strong linkage disequilibria for *C. albicans* (32, 61, 62, 64, 65).

The capacity to discriminate between highly related isolates, the basis for monitoring microevolution and substrain shuffling, was dissimilar between RAPD analysis, MLEE, analysis with Ca3, and analysis with CARE2. All methods did find the pairs of isolates FC-1 and FC-2 and isolates FC-3 and FC-4 to be identical (Table 3), and this is an expected result since each pair represented switch phenotypes of the same strain. However, the three methods showed different discriminatory capacities for the highly related but nonidentical pairs of isolates FC-11 and FC-12, FC-13 and FC-14, FC-17 and FC-18, FC-19 and FC-20, and FC-23 and FC-24 (Table 3). The RAPD method found two of these pairs to be identical and three to be highly related, while the MLEE method found three of these pairs to be identical, one pair to be highly related, and one pair to be moderately related (Table 3). Analysis with Ca3 found one of these pairs to be identical, two to be highly related, and two to be moderately related, while analysis with CARE2 found only one of these pairs to be identical, three to be highly related, and one to be moderately related (Table 3). The analyses with both Ca3 and CARE2 therefore found the same pair of isolates to be identical and the same two pairs of isolates to be highly related (Table 3). These latter two methods, which are based on the hybridization patterns generated with moderately repetitive sequences, are therefore more discriminatory in identifying microevolutionary changes between isolates with a common origin.

Linkage disequilibria evidenced within the RAPD analysis and MLEE data sets. The RAPD method involves amplification of unknown portions of the genome, and the MLEE method involves an analysis of defined genetic loci, presumably distinct from those amplified by the RAPD method. A consequence of sexual reproduction is random recombination between loci, and a consequence of clonal reproduction is the nonrandom association of loci (linkage disequilibrium). Both data sets were tested for nonrandom association between loci under the null hypothesis of random recombination (panmixia) for the 22 isolates of unrelated origin. The following four probability values were computed: *d1*, the combinatorial probability of sampling the most frequent genotype as often as or more often than actually observed if there were random recombination; *d2*, the probability of observing any genotype as often as or more often than the most common genotype actually observed if there were random recombination; *e*, the probability of observing as few or fewer genotypes than actually observed if there were random recombination; and *f*, the probability of observing as large or larger genetic disequilibrium than actually observed if there were random recombination (62, 65). If a probability is nonsignificant (5×10^{-2}), random recombination cannot be rejected, but if it is significant ($<5 \times 10^{-2}$), it supports the nonrandom association of loci. In all cases for the data obtained by both RAPD analysis and MLEE, the probabilities of a nonrandom association between loci were significant (Table 4), supporting the hypothesis that recombination is not random and that *C. albicans* reproduction is clonal.

TABLE 4. Linkage disequilibrium evidenced for RAPD and MLEE methods applied to *C. albicans*

Method	Level of significance[a]			
	d1	*d2*	*e*	*f*
RAPD analysis	8.8×10^{-5}	2.2×10^{-3}	$<1 \times 10^{-4}$	$<1 \times 10^{-4}$
MLEE	2×10^{-4}	4.5×10^{-2}	4.5×10^{-2}	$<1 \times 10^{-4}$

[a] Levels of significance for nonrandom association between loci (linkage disequilibrium) under the null hypothesis of random recombination. Tests were done with the sample of 22 isolates from unrelated hosts. *d2*, *e*, and *f* are based on Monte Carlo simulations with 10^4 iterations. See text for definitions.

DISCUSSION

With the development of a variety of DNA-based techniques for comparing independent isolates of the infectious yeast *C. albicans*, an ever increasing number of published studies have assessed relatedness between isolates in a collection in order to answer specific epidemiological questions. In most of these studies, however, no attempt is made to verify that the fingerprinting method is, in fact, discriminatory between genetically unrelated strains, that the method can identify the same strain in repeat isolates, and that the method can discriminate between highly related but nonidentical isolates. Unfortunately, most researchers interpret fingerprinting patterns qualitatively and rarely use available computer-assisted systems, even in relatively large epidemiological studies. Some fingerprinting systems provide only a few bands for discrimination, in which cases the efficacy of the method for large epidemiological studies is questionable. Some fingerprinting methods are based on patterns which have been shown to evolve at rates affected by the physiology or phenotype of the cell. For instance, electrophoretic karyotypes have been shown to change at highly elevated rates when cells are in a high-frequency mode of phenotypic switching (40). The patterns generated by these latter methods will not reflect genetic distance between strains within a species. An effective fingerprinting method must (i) discriminate between unrelated strains, (ii) identify the same strain or highly related strains in independent isolates, (iii) generate a pattern that is complex enough to provide a quantitative measure of genetic relatedness between two isolates (i.e., a meaningful similarity coefficient), (iv) generate a pattern amenable to automatic computer-assisted analysis, (v) be highly reproducible within a laboratory and between laboratories, (vi) be verified by an independent method of measuring genetic distances (i.e., an unrelated fingerprinting method), and (vii) reveal microevolutionary changes between isolates in an infecting population.

Parity of RAPD analysis, MLEE, and hybridization with Ca3. Each of the methods tested involves comparisons of polymorphisms, and we have assumed that the DNA sequences analyzed by the four methods are unrelated. There is no homology between Ca3 and CARE2 (20, 26), no indication that any of the primers result in repetitive sequence amplification, and no indication that any of the loci analyzed by MLEE are present in either the Ca3 probe or the CARE2 probe. It is also unlikely, simply on the basis of chance, that any of the sequences amplified by the eight primers represent the known genes analyzed by MLEE.

The patterns generated by each method for the collection of isolates were compared with the same similarity coefficient, and in each case dendrograms based upon the same unweighted pair group method were generated (55). Each method determined as identical switch phenotypes of the same strain and as identical, highly related or moderately related

isolates with related host origins. RAPD analysis, MLEE, and analysis with Ca3 grouped the majority of isolates of unrelated origin into the same three major clusters. The concordance of the results of RAPD analysis, MLEE, and analysis with Ca3 was not perfect. While all three methods grouped the same sets of isolates into clusters II and III, a few isolates in clusters I obtained by the RAPD analysis and MLEE fell outside of the threshold for cluster I in the analysis with Ca3. These particular isolates, however, were weakly associated with the clusters obtained by RAPD analysis or MLEE. Despite these few exceptions, the clustering capacities and cluster similarities of the three methods were remarkable.

The results of analysis with CARE2, however, did not exhibit the same level of concordance observed for the results of the three other methods, and this weakness was examined in a detailed comparison with the clustering capabilities of analysis with Ca3. Analysis with CARE2 grouped all but 1 of the 10 isolates grouped by analysis with Ca3 in cluster I and all but 3 of the 10 isolates grouped by analysis with Ca3 in cluster II. For the isolates not grouped in clusters I and II by analysis with CARE2, the links in the corresponding clusters in the dendrogram generated by analysis with Ca3 were relatively weak (i.e., just above the threshold for clustering). However, analysis with CARE2 was not as effective in clustering the isolates of unrelated origin in group III. Differences in the patterns obtained by analysis with CARE2 are based primarily on the reorganization of a single repeat element (20), while differences in the patterns obtained by analysis with Ca3 are based on changes in moderately variant as well as hypervariable sequences (1, 25). The pattern obtained by analysis with CARE2 for different isolates may therefore have a tendency to evolve into unrelated patterns at significantly faster rates than the pattern obtained by analysis with Ca3. Isolates in group III appear to be undergoing disproportionately faster rates of reorganization in their CARE2 sequences, resulting in the dismantling of cluster III in the dendrogram generated by analysis with CARE2. Analysis with CARE2 therefore appears to be a less effective fingerprinting procedure than the other three methods, especially for group III isolates.

Effectiveness of the four methods in determining microevolutionary changes. The Ca3 probe is complex (1) and has both the capacity to group isolates of unrelated origin and the capacity to identify microevolutionary changes within infecting clonal populations over time (25, 26, 49, 52). The latter discriminatory capacity is based on hypervariable sequences in the *C. albicans* genome identified by the C fragment of the probe (1, 25). One would therefore expect analyses with both Ca3 and CARE2, which generate patterns based exclusively on the variability of a repeat element, to have greater resolving power than either RAPD analysis or MLEE for microevolutionary changes within a strain, the kinds of changes examined here between isolates of related origin. Indeed, we have found that analysis with Ca3 has the best resolving power and that analysis with CARE2 has the second best resolving power for microevolutionary change, but we have also found that both RAPD analysis and MLEE have modest degrees of resolving power.

Evidence for strong linkage disequilibria. We earlier demonstrated that *C. albicans* isolates from a population of HIV-positive individuals in France had a clonal population structure (38), and more recently, Boerlin et al. (8) and Graser et al. (15) demonstrated an overall clonal population structure in *C. albicans* isolates from populations of HIV-positive individuals in Switzerland and healthy individuals in the United States (Duke University, Durham, N.C.), respectively. Here, we have demonstrated that there is a strong correlation between polymorphic markers assessed by the four independent fingerprinting methods and that the nonrandom association between loci assessed by RAPD analysis and MLEE provides further evidence for strong linkage disequilibria. The combined results of the four fingerprinting methods suggest that the majority of isolates subdivide into three genetic groups between which genetic exchange is either rare (15) or absent. The robustness of the clusters is further suggested by the identification of the synapomorphic *Pep-3A* allele in all members of cluster II. All cluster II isolates, including the weakly associated isolates FC-3 and FC-4, contained the *Pep-3A* allele. No isolates in clusters I and III and no unclustered isolates contained this allele. Cluster II isolates were obtained from individuals in Ann Arbor, Mich.; El Paso, Tex.; and Iowa City, Iowa. Therefore, these strains appear to have evolved from a common progenitor, and they appear to have been disseminated throughout the continental United States.

Effectiveness of the four fingerprinting methods tested. We have demonstrated that RAPD analysis, MLEE, and hybridization analyses with Ca3 and CARE2 are all effective methods for assessing the genetic relatedness of *C. albicans* isolates in large epidemiological studies. Analysis with CARE2 may be less effective in clustering less related strains. In the case of the RAPD analysis and MLEE, there are still a few caveats, especially when applying them to large-scale studies. The RAPD method proved to be the method that could be performed the most quickly, but it was the least reproducible for low-intensity bands. Introducing RAPD data into a data bank in order to compute S_{AB}s also proved to be slower than the probe methods. Performance of the MLEE method, although highly reproducible, proved to be the most time-consuming, and again, more time was required to introduce the data from MLEE into a data bank for the computation of S_{AB}s. Southern blot hybridization with the Ca3 and CARE2 probes proved to be highly reproducible and relatively fast, and hybridization with Ca3 proved to be the best indicator of microevolution. All fingerprinting systems were amenable to computer-assisted methods. It is quite likely that additional fingerprinting methods will be developed in the near future with all of the attributes given above plus higher resolution. Regardless of the fingerprinting method used for comparative analysis, investigators should assess and verify their system before applying it to large-scale epidemiological studies.

ACKNOWLEDGMENTS

This work was supported by Public Health Services grants AI2392 and DE10758 from the National Institutes of Health (to D.R.S.) and grant ACC-SV7 from the French Ministry of Research (to M.T.). S.R.L. was supported by training grant AG00214 from the National Institutes of Health.

We are indebted to John Hellstein (U.S. Army, El Paso, Tex.), Rudolph Galask (University of Iowa, Iowa City) and Barbara Reed (University of Michigan, Ann Arbor), who participated in the collection of the isolates used in this study.

REFERENCES

1. **Anderson, J., T. Srikantha, B. Morrow, S. H. Miyasaki, T. C. White, N. Agabian, J. Schmid, and D. R. Soll.** 1993. Characterization and partial nucleotide sequence of the DNA fingerprinting probe Ca3 of *Candida albicans*. J. Clin. Microbiol. **31:**1472–1480.
2. **Asakura, K., S. Iwaguchi, M. Homma, T. Sukai, K. Higashide, and K. Tanaka.** 1991. Electrophoretic karyotypes of clinically isolated yeasts of *Candida albicans* and *C. glabrata*. J. Gen. Microbiol. **137:**2531–2538.
3. **Barchiesi, F., R. J. Hollis, M. Del Poeta, D. A. McGough, G. Scalise, M. G. Rinaldi, and M. A. Pfaller.** 1995. Transmission of fluconazole-resistant *Candida albicans* between patients with AIDS and oropharyngeal candidiasis documented by pulse-field gel electrophoresis. Clin. Infect. Dis. **21:**561–564.
4. **Bart-Delabesse, E., P. Boiron, A. Carlotti, and B. Dupont.** 1993. *Candida albicans* genotyping in studies with patients with AIDS developing resistance to fluconazole. J. Clin. Microbiol. **31:**2933–2937.

5. **Bart-Delabesse, E., H. van Deventer, W. Goessens, J. L. Poirot, N. Lioret, A. van Belkum, and F. Dromer.** 1995. Contribution of molecular typing methods and antifungal susceptibility testing to the study of a candidemia cluster in a burn care unit. J. Clin. Microbiol. **33:**3278–3283.
6. **Barton, R. C., A. van Belkum, and S. Scherer.** 1995. Stability of karyotype in serial isolates of *Candida albicans* from neutropenic patients. J. Clin. Microbiol. **33:**794–796.
7. **Boerlin, P., F. Boerlin-Petzold, C. Durussel, M. Addo, J.-L. Pagani, J.-P. Chave, and J. Bille.** 1995. Cluster of oral atypical *Candida albicans* isolates in a group of human immunodeficiency virus-positive drug users. J. Clin. Microbiol. **33:**1129–1135.
8. **Boerlin, P., F. Boerlin-Petzold, J. Goudet, C. Durussel, J.-L. Pagani, J.-P. Chave, and J. Bille.** 1996. Typing *Candida albicans* oral isolates from human immunodeficiency virus-infected patients by multilocus enzyme electrophoresis and DNA fingerprinting. J. Clin. Microbiol. **34:**1235–1248.
9. **Bostock, A., M. N. Khattak, R. Matthews, and J. Burnie.** 1993. Comparison of PCR fingerprinting with other molecular typing methods for *Candida albicans*. J. Gen. Microbiol. **139:**2179–2184.
10. **Caugant, D. A., and P. Sandven.** 1993. Epidemiological analysis of *Candida albicans* strains by multilocus enzyme electrophoresis. J. Clin. Microbiol. **31:**215–220.
11. **Cutler, J. E., P. M. Glee, and H. L. Horn.** 1988. *Candida albicans* and *Candida stellatoidea* specific DNA fragment. J. Clin. Microbiol. **26:**1720–1724.
12. **Doi, M., M. Homma, A. Chindamporn, and K. Tanaka.** 1992. Estimation of chromosome number and size by pulsed-field gel electrophoresis (PFGE) in medically important *Candida* species. J. Gen. Microbiol. **138:**2241–2251.
13. **Doi, M., M. Homma, S. I. Iwaguchi, K. Horibe, and K. Tanaka.** 1994. Strain relatedness of *Candida albicans* strains isolated from children with leukemia and their bedside parent. J. Clin. Microbiol. **32:**2253–2259.
14. **Doi, M., I. Mizuguchi, M. Hamma, and K. Tanaka.** 1994. Electrophoretic karyotypes of *Candida* yeasts recurrently isolated from single patients. Microbiol. Immunol. **38:**19–23.
15. **Graser, Y., M. Volovsek, J. Arrington, G. Shönian, W. Presber, T. G. Mitchell, and R. Vilgalys.** 1996. Molecular markers reveal that population structure of the human pathogen *Candida albicans* exhibits both clonality and recombination. Proc. Natl. Acad. Sci. USA **93:**12473–12477.
16. **Hellstein, J., H. Vawter-Hugart, P. Fotos, J. Schmid, and D. R. Soll.** 1993. Genetic similarity and phenotypic diversity of commensal and pathogenic strains of *Candida albicans* isolated from the oral cavity. J. Clin. Microbiol. **31:**3190–3199.
17. **Holmberg, K., and F. Feroze.** 1996. Evaluation of an optimized system for random amplified polymorphic DNA (RAPD)-analysis for genotypic mapping of *Candida albicans* strains. J. Clin. Lab. Anal. **10:**59–69.
18. **Horowitz, B. J., S. W. Edelstein, and L. Lippman.** 1987. Sexual transmission of *Candida*. Obstet. Gynecol. **69:**883–886.
19. **Kleinegger, C. L., S. R. Lockhart, K. Vargas, and D. R. Soll.** 1996. Frequency, intensity, species, and strains of oral *Candida* vary as a function of host age. J. Clin. Microbiol. **34:**2246–2254.
20. **Lasker, B. A., L. S. Page, T. J. Lott, and G. S. Kobayashi.** 1992. Isolation, characterization, and sequencing of *Candida albicans* repetitive element 2. Gene **116:**51–57.
21. **Le Guennec, R., J. Reynes, M. Mallie, C. Pujol, F. Janbon, and J.-M. Bastide.** 1995. Fluconazole- and itraconazole-resistant *Candida albicans* strains from AIDS patients: multilocus enzyme electrophoresis analysis and antifungal susceptibilities. J. Clin. Microbiol. **33:**2732–2737.
22. **Lehmann, P. F., B. J. Kemker, C.-B. Hsiao, and S. Dev.** 1989. Isoenzyme biotypes of *Candida* species. J. Clin. Microbiol. **27:**2514–2521.
23. **Lehmann, P. F., D. Lin, and B. A. Lasker.** 1992. Genotypic identification and characterization of species and strains within the genus *Candida* by using random amplified polymorphic DNA. J. Clin. Microbiol. **30:**3249–3254.
24. **Lischewski, A., M. Ruhnke, I. Tennagen, G. Schönian, J. Morschhauser, and J. Hacker.** 1995. Molecular epidemiology of *Candida* isolates from AIDS patients showing different fluconazole resistance profiles. J. Clin. Microbiol. **33:**769–771.
25. **Lockhart, S. R., J. J. Fritch, A. S. Meier, K. Schröppel, T. Srikantha, R. Galask, and D. R. Soll.** 1995. Colonizing populations of *Candida albicans* are clonal in origin but undergo microevolution through C1 fragment reorganization as demonstrated by DNA fingerprinting and C1 sequencing. J. Clin. Microbiol. **33:**1501–1509.
26. **Lockhart, S. R., B. D. Reed, C. L. Pierson, and D. R. Soll.** 1996. Most frequent scenario for recurrent *Candida* vaginitis is strain maintenance with "substrain shuffling": demonstration by sequential DNA fingerprinting with probes Ca3, C1, and CARE2. J. Clin. Microbiol. **34:**767–777.
27. **Lupetti, A., G. Guzzi, A. Pladini, K. Swart, M. Campa, and S. Senesi.** 1995. Molecular typing of *Candida albicans* in oral candidiasis: karyotype epidemiology with human immunodeficiency virus-seropositive patients in comparison with that with healthy carriers. J. Clin. Microbiol. **33:**1238–1242.
28. **Magee, P. T., L. Bowdin, and J. Staudinger.** 1992. Comparison of molecular typing methods for *Candida albicans*. J. Clin. Microbiol. **30:**2674–2679.
29. **Mahrous, M., T. J. Lott, S. A. Meyer, A. D. Awant, and D. G. Ahearn.** 1990. Electrophoretic karyotyping of typical and atypical *Candida albicans*. J. Clin. Microbiol. **28:**876–881.
30. **Mantel, N.** 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. **27:**209–220.
31. **Merz, W. G., C. Connelly, and P. Hieter.** 1988. Variation of electrophoretic karyotypes among clinical isolates of *Candida albicans*. J. Clin. Microbiol. **26:**842–845.
32. **Miller, R. D., and D. L. Hartl.** 1987. Biotyping confirms a nearly clonal population structure in *Escherichia coli*. Evolution **40:**1–12.
33. **Millon, L., A. Manteaux, G. Reboux, C. Drobacheff, M. Monod, T. Barale, and Y. Michel-Briand.** 1994. Fluconazole-resistant recurrent oral candidiasis in human immunodeficiency virus-positive patients: persistence of *Candida albicans* strains with the same genotype. J. Clin. Microbiol. **32:**1115–1118.
34. **Ng, T. T., and D. W. Denning.** 1993. Fluconazole resistance in *Candida* in patients with AIDS—a therapeutic approach. J. Infect. **26:**117–125.
35. **Odds, F. C.** 1987. *Candida* infections: an overview. Crit. Rev. Microbiol. **15:**1–5.
36. **Odds, F. C.** 1988. *Candida* and candidosis: a review and bibliography, 2nd ed. Balliere Tindall, London, United Kingdom.
37. **Odds, F. C., and A. B. Abbott.** 1980. A simple system for the presumptive identification of *Candida albicans* and differentiation of strains within the species. Sabouraudia **18:**301–317.
38. **Pujol, C., J. Reynes, F. Renaud, M. Raymond, M. Tibayrenc, F. J. Ayala, F. Janbon, M. Mallie, and J.-M. Bastide.** 1993. The yeast *Candida albicans* has a clonal mode of reproduction in a population of infected human immunodeficiency virus-positive patients. Proc. Natl. Acad. Sci. USA **90:**9456–9459.
39. **Pujol, C., J. Reynes, F. Renaud, M. Mallié, and J.-M. Bastide.** 1993. Analyse génétique de souches de *Candida albicans* par électrophorèse des isoenzymes. J. Mycol. Med. **3:**14–19.
40. **Ramsey, H., B. Morrow, and D. R. Soll.** 1994. An increase in switching frequency correlates with an increase in recombination of the ribosomal chromosomes of *Candida albicans* strain 3153A. Microbiology **140:**1525–1531.
41. **Reynes, J., C. Pujol, C. Moreau, M. Mallie, F. Renaud, F. Janbon, and J.-M. Bastide.** 1996. Simultaneous carriage of *Candida albicans* strains from HIV-infected patients with oral candidiasis: multilocus enzyme electrophoresis analysis. FEMS Microbiol. Lett. **137:**269–273.
42. **Robert, F., F. Lebreton, M. E. Bougnoux, A. Paugam, D. Wassermann, M. Schlotterer, S. C. Tourte, and C. J. Dupouy.** 1995. Use of random amplified polymorphic DNA as a typing method for *Candida albicans* in epidemiological surveillance of a burn unit. J. Clin. Microbiol. **33:**2366–2371.
43. **Sadhu, C., M. J. McEachern, B. E. Rustchenko, J. Schmid, D. R. Soll, and J. B. Hicks.** 1991. Telomeric and dispersed repeat sequences in *Candida* yeasts and their use in strain identification. J. Bacteriol. **173:**842–850.
44. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
45. **Scherer, S., and D. A. Stevens.** 1987. Application of DNA typing methods to epidemiology and taxonomy of *Candida* species. J. Clin. Microbiol. **25:**675–679.
46. **Scherer, S., and D. A. Stevens.** 1988. A *Candida albicans* dispersed, repeated gene family and its epidemiologic applications. Proc. Natl. Acad. Sci. USA **85:**1452–1456.
47. **Schmid, J., E. Voss, and D. R. Soll.** 1990. Computer-assisted methods for assessing strain relatedness in *Candida albicans* by fingerprinting with the moderately repetitive sequence Ca3. J. Clin. Microbiol. **28:**1236–1243.
48. **Schmid, J., F. C. Odds, M. J. Wiselka, K. G. Nicholson, and D. R. Soll.** 1992. Genetic similarity and maintenance of *Candida albicans* strains from a group of AIDS patients, demonstrated by DNA fingerprinting. J. Clin. Microbiol. **30:**935–941.
49. **Schmid, J., M. Rotman, B. Reed, C. L. Pierson, and D. R. Soll.** 1993. Genetic similarity of *Candida albicans* strains from vaginitis patients and their partners. J. Clin. Microbiol. **31:**39–46.
50. **Schmid, J., Y. P. Tay, L. Wan, M. Carr, D. Parr, and W. McKinney.** 1995. Evidence for nosocomial transmission of *Candida albicans* obtained by Ca3 fingerprinting. J. Clin. Microbiol. **33:**1223–1230.
51. **Schönian, G., O. Meusel, H.-J. Tietz, W. Meyer, Y. Gräser, I. Tausch, W. Presber, and T. G. Mitchell.** 1993. Identification of clinical strains of *Candida albicans* by DNA fingerprinting with the polymerase chain reaction. Mycoses **36:**171–179.
52. **Schröppel, K., M. Rotman, R. Galask, K. Mac, and D. R. Soll.** 1994. Evolution and replacement of *Candida albicans* strains during recurrent vaginitis demonstrated by DNA fingerprinting. J. Clin. Microbiol. **32:**2646–2654.
53. **Slutsky, B., J. Buffo, and D. R. Soll.** 1985. High-frequency switching of colony morphology in *Candida albicans*. Science **230:**666–669.
54. **Slutsky, B., M. Staebell, J. Anderson, L. Risen, M. Pfaller, and D. R. Soll.** 1987. "White-opaque transition": a second high-frequency switching system in *Candida albicans*. J. Bacteriol. **169:**189–197.
55. **Sneath, P. H., and R. R. Sokal.** 1973. Numerical taxonomy. The principles and practice of numerical classification, p. 230–234. W. H. Freeman & Co., San Francisco, Calif.

56. **Soll, D. R.** 1993. DNA fingerprinting of *Candida albicans*. J. Mycol. Med. **3:**37–44.
57. **Soll, D. R., C. J. Langtimm, J. McDowell, J. Hicks, and R. Galask.** 1987. High-frequency switching in *Candida* strains isolated from vaginitis patients. J. Clin. Microbiol. **25:**1611–1622.
58. **Soll, D. R., M. Staebell, C. Langtimm, M. Pfaller, J. Hicks, and T. V. Rao.** 1988. Multiple *Candida* strains in the course of a single systemic infection. J. Clin. Microbiol. **26:**1448–1459.
59. **Soll, D. R., R. Galask, J. Schmid, C. Hanna, K. Mac, and B. Morrow.** 1991. Genetic dissimilarity of commensal strains of *Candida* spp. carried in different anatomical locations of the same healthy women. J. Clin. Microbiol. **29:**1702–1710.
60. **Sullivan, D., D. Bennett, M. Henman, P. Harwood, S. Flint, F. Mulcahy, D. Shanley, and D. Coleman.** 1993. Oligonucleotide fingerprinting of isolates of *Candida* species other than *C. albicans* and of atypical *Candida* species from human immunodeficiency virus-positive and AIDS patients. J. Clin. Microbiol. **31:**2124–2133.
61. **Tibayrenc, M., and F. J. Ayala.** 1988. Isozyme variability of *Trypanosoma cruzei*, the agent of Chagas disease: genetical, taxonomical and epidemiological significance. Evolution **42:**277–292.
62. **Tibayrenc, M., F. Kjellberg, and F. J. Ayala.** 1990. A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. Proc. Natl. Acad. Sci. USA **87:**2414–2418.
63. **Tibayrenc, M., F. Kjellberg, J. Arnaud, B. Oury, F. Brenière, M.-L. Dardé, and F. J. Ayala.** 1991. Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. Proc. Natl. Acad. Sci. USA **88:**5129–5133.
64. **Tibayrenc, M., K. Neubauer, C. Barnabé, F. Guerrini, D. Skarecky, and F. J. Ayala.** 1993. Genetic characterization of six parasitic protozoa: parity between random-primer DNA typing and multilocus enzyme electrophoresis. Proc. Natl. Acad. Sci. USA **90:**1335–1339.
65. **Tibayrenc, M.** 1995. Population genetics of parasitic protozoa and other microorganisms. Adv. Parasitol. **36:**47–115.
66. **Vazquez, J. A., A. Beckley, J. D. Sobel, and M. J. Zervos.** 1991. Comparison of restriction enzyme analysis versus pulse-field gradient gel electrophoresis as a typing system for *Candida albicans*. J. Clin. Microbiol. **29:**962–967.
67. **Vazquez, J. A., J. D. Sobel, R. Demitriou, J. Vaishampayan, M. Lynch, and M. Zervos.** 1994. Karyotyping of *Candida albicans* isolates obtained longitudinally in women with recurrent vulvovaginal candidiasis. J. Infect. Dis. **170:**1566–1569.
68. **Venkateswarlu, K., D. W. Denning, N. J. Manning, and S. L. Kelly.** 1995. Resistance to fluconazole in *Candida albicans* from AIDS patients correlated with reduced intracellular accumulation of drug. FEMS Microbiol. Lett. **131:**337–341.
69. **Voss, A., M. A. Pfaller, R. J. Hollis, J. Rhine-Chalberg, and B. N. Doebbeling.** 1995. Investigation of *Candida albicans* transmission in a surgical intensive care unit cluster by using genomic DNA typing methods. J. Clin. Microbiol. **33:**576–580.
70. **Whelan, W. L., D. R. Kirsch, K. J. Kwon-Chung, S. M. Wahl, and P. D. Smith.** 1990. *Candida albicans* in patients with the acquired immunodeficiency syndrome: absence of a novel or hypervirulent strain. J. Infect. Dis. **162:**513–518.